



LUNDS
UNIVERSITET

INSTITUTIONEN FÖR PSYKOLOGI

Förmågan till röstigenkänning avseende expertis, konfidens
och verbal överskuggning

Charlotta Sjösten
Christofer Linder

Kandidatuppsats VT 2011

Handledare: Mikael Johansson
Susanna Bernstrup

Abstract

Syftet med studien var att utreda om en expertgrupp, bestående av 24 deltagare rekryterade från en larmcentral, hade en bättre förmåga att minnas röster än en kontrollgrupp bestående av 24 deltagare. Försöksdeltagarna testades i ett experiment bestående av en röstdel och en ansiktsdel, varje del bestod av en verbaliseringsdel, som var tänkt att testa fenomenet verbal överskuggning samt en icke-verbaliseringsdel. Det förväntades att expertgruppen skulle prestera bättre i röstdelen av experimentet på grund av deras yrkeserfarenhet, samt ha lägre konfidens än kontrollgruppen. Resultatet från studien visade att expertgruppen inte var bättre än kontrollgruppen på att minnas röster. Dock hade expertgruppen en signifikant högre konfidens avseende sin förmåga att minnas röster. Resultatet kan ha berott på att expertgruppens erfarenhet skapade en övertro på deras minnesförmåga istället för en försiktighet. Studien visade på att det är svårare att minnas röster än ansikte samt att försiktighet bör iaktas angående konfidens och korrekthet, även för experter.

Nyckelord: Röstigenkänning, expertis, konfidens, verbal överskuggning, vittne.

Introduktion

Det är en sen fredagskväll, du kommer gående längs med grusgången i stadsparken. Du har låtit tankarna flyga iväg till något trevligare än det höstrusk du just nu skyndar dig fram igenom. Från ingenstans kliver en mörkklädd figur fram med en huva neddragen framför ansiktet. Du ser en kniv blänka i mörkret och hör en röst säga ”hit med plånboken och mobilen”. Det hela är över på mindre än en minut.

Det ringer in till vaktkuren på polisstationen på Davidhall i Malmö, den sömnige skyddsvakten svarar och får höra en lugn röst som meddelar att ”det ligger en väska med en bomb i på Stortorget”.

TVå till synes helt skilda händelser, det ena ett enkelt personrån som händer varje dag det andra ett potentiellt terrorhot. Men ändå är de ganska lika, i båda fallen är det gärningsmannens röst som inledningsvis ett av de få spår som finns att jobba utifrån. Och som i de flesta fall, i alla fall när det gäller personrån, våldtäkter och kränkande telefonsamtal finns gärningsmannens röst inte inspelad. Oftast finns det teknisk utrustning som bandar alla samtal på polisens ledningscentraler, SOS-centraler m.m. men teknisk utrustning kan fallera. I de lägena finns bara brottsoffret och eventuella vittnens minne att gå på när det gäller att identifiera gärningsmannens röst.

Denna studie är tänkt att titta närmare på människors förmåga att minnas röster, eller som de kallas i brottssammanhang, öronvittnen. Fokus i studien kommer att ligga på att utreda om det finns grupper som är bättre än andra på att minnas röster, exempelvis om vissa yrkesgrupper förvärvar eller utvecklar sådana kunskaper genom sitt arbete.

Öronvittnen

Oftast när det pratas om vittnen i media eller i forensiska sammanhang kan det lätt tas för givet att det pratas om ett vittne som har sett något med sina ögon, ett så kallat ögonvittne. Detta gör att termen öronvittne känns lite obekant. Enligt Yarmey (1995) finns det otaliga empiriska studier som undersöker människans förmåga att minnas vad hon har sett. Människors förmåga att minnas vad hon har hört är emellertid ett forskningsområde som inte alls är lika undersökt (Yarmey, 1995). De exempel som inledningsvis gavs pekar på att en hel del brott inträffar utan att det finns goda möjligheter för vittnet att kunna göra ett visuellt utpekande. Detta tydliggör att öronvittnen är viktiga och att mer kunskap om hur minnet för röster fungerar är nödvändigt (Clifford & Bull, 1978; Clifford, 1980; Yarmey, 1995). En ökad kunskap inom detta område bidrar både till att lättare kunna veta om, och under vilka

förhållande, ett vittnes utsaga eller utpekande av en röst är tillförlitligt. Men mer forskning inom området kan också bidra till att bättre förhållningsstrategier och tekniker kan utvecklas, vilket kan bidra till bättre och mer tillförlitliga vittnesutsagor.

Bakgrund

Ett klassiskt exempel på ett utpekande som skett med hjälp av öronvittnen är kidnappningen av Charles Lindberghs barn 1932. Den berömde flygaren Charles Lindberghs 1,5 år gamla son blev kidnappad och hittades sedan mördad ett par månader senare. Kidnapparna hade lämnat en lapp där de krävde en lösensumma på 50 000 dollar. I ett försök att överlämna lösensumman befann sig sedan Charles Lindbergh senare på kvällen på en kyrkogård där han hörde en man ropa "*Hey doctor! Here doctor, over here*". Mannen som uttalat orden hade brutit på tyska. Och efter två års utredning så greps en tysk invandrare vid namn Bruno Richard Hauptmann. Lindbergh pekade sedan ut Hauptmann som den röst han hade hört på kyrkogården mer än två år tidigare. Hauptmann dömdes till döden till stor del på grund av Lindberghs utpekande, det fanns inte mycket annan bevisning mot Hauptmann. (Granhag, 2001; Read & Craik, 1995; Solan & Tiersma, 2003).

Frågan om det var möjligt för en person att göra ett utpekande om rösten efter mer än två år väckte många funderingar. Den uppmärksamhet som Lindbergh-fallet skapade bidrog till att det började forskas inom öronvittnesområdet. Den första forskningen som genomfördes inom detta område gjordes av Francis McGehee, som var intresserad av just Lindbergh-fallet. McGehee var tveksam till möjligheten att bevara minnet av en röst under väldigt lång tid, så som skedde i Lindbergh-fallet. År 1937 började McGehee att testa hur minnet påverkades över tid, detta gjorde hon genom att dela in försöksdeltagarna i olika grupper beroende på hur lång tid som förflöt mellan den första exponeringen för rösten och utpekandet. Tiden kunde variera allt från 1 dag till 5 månader. McGehee kom fram till att upp till en vecka efter exponeringstillfället så hade försöksdeltagarna en ganska god förmåga att kunna peka ut vilken röst de hört, därefter skedde en märkbar försämring (Clifford & Bull, 1978; Granhag & Christianson, 2008).

Rösten

Eftersom företeelsen med öronvittnen och forskningen kring dem är ganska okänd enligt Yarmey (1995) kan det vara viktigt att inledningsvis i en studie klargöra vissa begrepp som är av vikt i den här typen av forskning, såsom exempelvis språkframställning. Tidigare forskning inom ämnet öronvittnen har enligt Yarmey (1995) inte tittat så mycket på de

fysiologiska eller neurologiska grunderna bakom tal. Fokus har istället legat på att ta reda på vilka möjligheter gemene man har att känna igen vissa grundläggande element hos en röst de har. Dessa element hos en röst är, tonhöjd, röststyrka och varaktighet. Yarmey (1995) jämför dessa tre element hos en röst med begrepp som används inom vanlig akustik, när det handlar om att beskriva andra sorters ljud. Tonhöjden har samband med frekvens, det som uttrycks i måttet Hertz, hur många gånger ljudet svänger under en sekund. Röststyrkan kan kopplas till decibelnivån, det vill säga hur stora svängningar ljudet skapar. Varaktigheten (millisekunder) kan ses som ett mått på längd (Yarmey 1995).

Män och kvinnors röster skiljer sig åt framförallt när det handlar tonhöjden. Oftast framstår en mansröst som mycket mörkare, eller djupare än en kvinnas röst. Detta beror på att en kvinnas stämband generellt sett är kortare och lättare än en mans, vilket gör att de kan vibrera fler gånger per sekund. Detta får till följd att en mansröst framstår som mörkare (Yarmey 1995). Hos en kvinna svänger stämbanden ca 200-400 gånger per sekund, vilket är samma sak som 200-400 hertz. Hos män är det istället 100-200 svängningar per sekund (Engstrand 2004).

Givetvis är det inte bara dessa tre sakerna, tonhöjd, röststyrka och längd, som påverkar hur människan låter. Vår röst påverkas även av sådana saker som regionala, kulturella och sociala skillnader, ett bra exempel på detta är vår dialekt (Yarmey, 1995).

Minnet

Självklart är det inte bara rösten som är viktig i denna studie, utan den måste sättas i samband med det mänskliga minnet och dess funktion. Minnet brukar delas upp i en rad kategorier, beroende på vilken funktion det fyller. Det semantiska minnet är där all vår kunskap lagras, exempelvis årtal, städer etc. Vidare finns proceduralt minnet, som kan ses som ett minne för handlingar, detta minne är mer implicit men där lagras vårt minne för vissa handlingar exempelvis att köra en bil eller att cykla. När det kommer till vittnespsykologi och vittnesmål i stort är det framförallt det episodiska minnet som är intressant. Det episodiska minnet är det minnet där våra personliga minnen lagras, exempelvis minnet av vad som hände förra julen eller vad som hände förra gången man gick genom stadsparken (Granhag, 2001; Granhag & Christianson, 2008).

Eftersom det episodiska minnet baserar sig på händelser som är självupplevda är det till stor del det partiet av minnet som behöver användas i vittnessammanhang. Minnet av hur den eventuella gärningsmannen såg ut, vad han/hon hade på sig, vad gärningsmannen sade etc. (Granhag, 2001; Granhag & Christianson, 2008). Minnet av en händelse är emellertid inte

konstant utan kan påverkas och förändras. Minnesprocessen för ett minne brukar delas in i tre faser, *inkodning*, *lagring* och *framplockning*. Under dessa tre processer kan minnet påverkas och förvrängas på olika sätt (Granhag & Christianson, 2008).

Under den minnesprocess som kallas inkodning så kodas minnet in. Under inkodningen kan olika omständigheter göra att minnet kodas in felaktigt eller bristfälligt. Stress är ett typiskt exempel på en faktor som kan påverka minnet negativt. Exempelvis kan stressen som uppkommer under en traumatisk eller hotfullsituation påverka. Det som händer då är att uppmärksamheten oftast skiftar till de elementen i situationen som framstår som mest hotfulla. Vapenfokuseringseffekten är ett bra exempel på detta, att offret fokuserar på exempelvis kniven gärningsmannen håller i sin hand eftersom det är där det hotfulla finns. Eftersom vittnets fokus läggs på kniven så kan detta få till följd att andra detaljer inte blir så väl inkodade i minnet, till exempel gärningsmannens klädsel, eller hur gärningsmannen röst lät. Vilket får till följd att minnet som skapas av händelsen påverkas negativt och det kan vara svårt att återge en helt korrekt bild (Granhag & Christianson, 2008). Det behöver emellertid inte vara så dramatiskt som vid ett brott utan det kan som Schacter (2001) nämner handla om att personen är ofokuserad eller tankspridd vid tidpunkten när minnet skapades vilket kan få till följd att minnet kodas in felaktigt eller bristfälligt.

Även under den minnesprocess som kallas lagring kan minnet förändras. Beroende på hur väl minnet kodats in kan minnet blekna över tid. Även om det inte försvinner helt så kan detaljer och kunskap försvinna om det går lång tid från inkodning till framplockning (Schacter, 2001). Minnet påverkas även av att en person diskuterar händelser kopplade till ett minne, vilket gör att minnet kan påverkas och förändras till något annat än hur det såg ut från början (Granhag & Christianson, 2008).

Den tredje fasen i minnesprocessen är den som kallas för framplockning. Det är då personen försöker plocka fram minnet för att exempelvis kunna berätta om det. Precis som vid de två andra faserna kan minnet påverkas och förvrängas. När någon blivit utsatt för ett brott och sedan ska förhöras är det framplockning det handlar om. I den situationen som skapas när en förhørsledare försöker hjälpa ett vittne eller brottsoffer att berätta vad som har hänt, kan dennes minne av händelsen påverkas. Det kan bland annat handla om att förhørsledaren ställer så kallade ledande frågor ”hur snabbt körde den blåa bilen”, när det möjligen var så att bilen i fråga var röd. Vittnet kan i det här läget omedvetet eller medvetet börja tvivla på sitt eget minne och börja tro att bilen faktiskt var blå, vilket innebär att det ursprungliga minnet kan förändras. Därför är det väldigt viktigt för en eventuell förhørsledare att vara på sin vakt och inte ställa några ledande frågor som riskerar att påverka vittnet, utan

hellre försöka be dem att berätta fritt själva. Detta är särskilt problematiskt när det gäller yngre barn som förhörs då de kan ha svårare än vuxna att berätta självständigt (Schacter, 2001; Granhag, 2001; Granhag & Christianson, 2008).

En intressant frågeställning, som kan kopplas både till den allmänna kunskapen om minnet samt vara relevant för denna studie, är den som Rämä och Courtney (2004) ställde. De ville titta närmare på är vilka delar av hjärnan som användes beroende av vilken typ av information som försöksdeltagarna skulle försöka att minnas. Försöksdeltagarna fick genomgå ett minnestest som var uppdelat i olika delar, en del där de fick ansikten presenterade för sig, och en annan del där de fick röster som de skulle försöka att minnas och kunna peka ut i ett senare skede. Det intressanta med denna studie var att försöksdeltagarna samtidigt låg i en fMRI-maskin, det vill säga en maskin som med hjälp av magnetkamera kan ta bilder av hjärnan och vilka delar av hjärnan som används. Syftet med studien var att se vilka områden i hjärnan som aktiverades vid respektive uppgift, samt att se om det var så att samma område i hjärnan användes för båda uppgifterna. Studien visade att den delen av hjärnan som kallas för fusiform gyrus (FG) och lateral occipital gyrus (LOG) var mer aktiverade när försöksdeltagaren skulle försöka minnas ansikten. När minnet för rösterna testades var det istället mer aktivering i de delar av hjärnan som kallas för superior temporal sulcus och gyrus (STS/STG). Vid första anblicken verkade det som att aktiveringen var lika stor i den delen av hjärnan som är benämnd som ventral prefrontal cortex, men vid närmare granskning visade det sig att aktiveringen skedde på olika platser beroende på om det handlade om röster eller ansikte (Rämä & Courtney, 2004).

Den uppdelning av aktivering i hjärnan beroende på vilken uppgift som presenterades väcker en hel del intressanta frågeställningar. Om det är så att det finns människor som överlag klarar den ena uppgiften bättre än den andra, som kan räknas som experter på den uppgiften så behöver de inte nödvändigtvis räknas som experter på den andra.

Verbal överskuggning

Verbal överskuggning är ett fenomen som kan störa det mänskliga minnet och därmed försvåra att korrekt vittnesmål ges. Verbal överskuggning beskriver fenomenet i vilken verbalisering negativt påverkar vår igenkänningsförmåga och vårt minne att senare behandla samma material. Det är själva verbaliseringen av något huvudsakligen icke verbalt stimuli, såsom ett ansikte eller en röst som utgör själva störningen (Dodson, Johnson & Schooler, 1997). Det finns tre kategorier som tillsammans eller var och en kan förklara fenomenet verbal överskuggning (Chin & Schooler, 2008). Dessa tre kategorier är *innehåll*, *processande*

samt *kriterium*. Processande och kriterium teorierna handlar om *sättet* vi tänker på, medan innehåll handlar om *vad* en tänker på (Chin & Schooler, 2008). *Innehåll* hänvisar till uppfattningen att det specifika innehållet i verbaliseringen stör originalminnet, vilket gör att prestationen blir sämre vid igenkänning. The recoding interference explanation ("omkodningsstörningsförklaring") föreslår att minnet för originalstimulis förstörs därför att försökspersonen förvirrar det verbala minnet skapat av det visuella stimulit med originalminnet (Chin & Schooler, 2008). Det är särskilt troligt att en verbal överskuggningseffekt sker för en person med högre perceptuella kunskaper än verbal förmåga. Melcher och Schooler (1996) bevisade detta genom att genomföra en vintestning med tre olika grupper, icke vindrickare, professionella vindrickare samt icke professionella vindrickare. De fann att det är störst risk för en icke professionell vindrickare att drabbas av verbal överskuggning, jämfört med en professionell vindrickare som har hög perceptuell och verbal kunskap eller jämfört med en icke vindrickare som saknar perceptuell och verbal kunskap angående vin. Den icke professionella vindrickaren har högre perceptuell kunskap om vin än verbal förmåga att beskriva det, så när den ska förklara vinet inträffar en verbal överskuggningseffekt till följd av en mismatch mellan innehållet i verbaliseringen och perceptionen av vinet. Denna slutsats stödjer teorin att just innehållet i verbaliseringen avgör om en verbal överskuggning sker (Melcher & Schooler, 1996). Innehållet i instruktionen experimentledaren ger avgör också huruvida verbal överskuggning överhuvudtaget sker och i vilken mån. Kräver testledaren att deltagarna ska göra en forcerad återkallelse, exempelvis att man blir tvingad att skriva 25 rader om något även om man endast känner sig behöva 5 rader, blir verbal överskuggning än starkare. Detta till följd av att forcerad återkallelse innebär att försökspersonen i den pressade situationen använder mer felaktig information som sedan stör igenkänningen av stimulit (Chin & Schooler, 2008).

Processande hänvisar till processandet testdeltagare utsätts för när de exempelvis uppmanas beskriva ett ansikte. Processen att gå från att se ett ansikte som ett, till att beskriva dess olika delar, en övergång från ett holistiskt synsätt till ett mer uppdelat. Detta processande kan vara en orsak till att verbal överskuggning sker när testdeltagaren sen ser det hela ansiktet igen (Chin & Schooler, 2008).

Kriterium hänvisar till att verbalisering tycks göra att personer väljer mer försiktigt. Försöksdeltagare undviker att peka ut någon från en lineup om de kan. Om en lineup möjliggör att välja "inga alternativ" korrekt kommer de som verbaliserat i högre utsträckning än de som inte verbaliserat att välja detta alternativ. Verbalisering förfinar och diskriminerar därmed responskriteriet. Om målindividens är närvarande i lineupsen kommer dock personer

som verbaliserat att uppvisa sämre prestanda, en överskuggningseffekt har skett (Chin & Schooler, 2008). Ursprungligen visats i ramen för ansiktsgenkänning (Schooler & Engstler-Schooler, 1990) har fenomenet med verbal överskuggning visat sig vara ganska generellt och verbal överskuggning har bevisats påverka för så vitt skilda områden som beslutsfattande, problemlösning, visuella imaginära bilder, mentala rotationer etc. (Chin & Schooler, 2008). Flertalet studier har visat på en verbal överskuggning för ansikten. De senaste åren har emellertid ett fåtal studier gjorts kring verbal överskuggning för röster. Vanags, Carroll och Perfect (2005) genomförde en studie där testdeltagarna fick lyssna till inspelade röster och sedan från en sexpersoners lineup identifiera den korrekta rösten. Verbal överskuggning påvisades, då den grupp som verbaliserat hade klart sämre resultat än kontrollgruppen som inte verbaliserat. Även Mitchell och MacDonald (2011) bevisade en verbal överskuggningseffekt i en studie som handlade om att känna igen en sångröst. Ett resultat från deras studie var även att konfidensen inte påverkades av att svaren var felaktiga.

Faktorer som påverkar ett öronvittne

Som Yarmey (1995) nämnde är det inte bara hur en röst låter som kan påverka minnet, utan även andra faktorer. Faktorer som kan påverka ett öronvittne kan delas upp i tre kategorier; situationsspecifika, röstspecifika samt vittnesspecifika, enligt Yarmey (1995).

Situationsspecifika faktorer

Situationsspecifika är relaterade till sammanhanget i vilket vittnet tar del av rösterna. Om det är en eller flera röster, ju flera röster desto svårare att känna igen. Har gärningsmännen pratat med varandra har vittnen svårare att identifiera en röst. Gärningsmannens tonfall, om gärningsmannen skriker vid tillfället och sedan vid en vittneskonfrontation pratar normalt kommer det vara svårare att känna igen rösten (Granhag, 2001).

Initial exponeringstid. Endast ett par sekunders exponeringstid försämrar vittnes igenkänning avsevärt. Ca 20 sek, som de flesta test är utformade stämmer väl överens med verkligheten. Dock uppmärksammar Granhag (2001) att desto längre exponeringstid, desto fler felaktiga identifikationer. Detta följer av att vittnen som exponerats för en röst en längre tid sällan avhåller sig från att göra en identifikation, till skillnad mot de som hört en röst kortare tid som inte känner sig säkra på att göra en identifikation och därmed avstår.

Planerad eller oplanerad inkodning. Vid studier förvarnas vittnen att röstigenkänning kommer ske, detta är ju inte hur det ser ut i verkligheten. Förvarnade öronvittnen är därmed bättre på att komma ihåg röster än icke förvarnade (Granhag, 2001; Yarmey 1995).

Röstspecifika faktorer

Bekant –Obekant röst. Forskning visar att desto mer bekant en röst är, desto lättare är det att identifiera den. Emellertid är en röst också förknippad med miljön man vanligen hör den i. Det är svårt att känna igen en röst som vanligen hörs i exempelvis skolan, om man sedan hör den på gymmet. I experiment utsätts de medverkande för röster de aldrig hört innan, därmed obekanta röster. Detta påverkar då enligt Granhag (2001) igenkänningen.

Förställd eller naturlig röst. Även små förvrängningar i rösten kan försämra chansen för ett vittne att känna igen en röst. Gärningsmannen har minst två chanser att förställa sin röst dels när öronvittnet hör den första gången samt när gärningsmannen talar in sin röst till röstkonfrontation. Att viska är ett framgångsrikt sätt att förställa rösten. Detta beror på att en viskad röst inte har tonfallsvariationer (Granhag, 2001). Enligt Yarmey, Yarmey, Yarmey och Parliament (2001) ska ett utpekande från ett vittne som hört en viskande röst tas emot väldigt försiktigt, eftersom det i deras studie visade sig att en röst som viskar är väldigt svår att komma ihåg vid ett senare tillfälle, även om vittnet har hört rösten under en ganska lång tid.

Accent. Dialekt och brytning är detaljer som påverkar vittnens tillförlitlighet, emellertid näst intill utforskat. Forskning har dock visat att öronvittnen har lättare att känna igen röster talade på sitt modersmål än på ett främmande språk (Granhag, 2001).

Enligt Mullennix et al. (2010) har vittnen en tendens att bedöma en mycket ljus röst som mycket ljusare än den egentligen var och vice versa för mörka röster vid tiden för ett utpekande. Detta jämfört med röster som varken räknades som ljusa eller mörka. Mullennix et al. (2010) menar att detta kan förklara varför det generellt sett är så svårt att minnas röster.

Vitnessspecifika faktorer

Vitnessspecifika faktorer som relaterar, enligt Yarmey (1995) och Granhag (2001), till vittnens förmåga att minnas röstbaserad information är följande:

Öronvittnets ålder. Äldre människor har generellt sämre hörsselförmåga. Unga barn har ett inte fullt utvecklat sinne för att känna igen röster. Öronvittnens hörsel bör förstås generellt iakttas.

Öronvittnes kön. Studier från -30-40 talet visar på att män kan identifiera manliga röster bättre och kvinnor kan identifiera kvinnoröster bättre. Modern forskning har dock inte bevisat detta (Granhag, 2001).

Tränat, otränat öronvittne. Det är svårt att träna sig till att bli ett bättre öronvittne. Emellertid är det förstås möjligt att öva upp sin uppmärksamhetsförmåga. En missuppfattning kan vara att blinda öronvittnen skulle ha bättre förmåga att lyssna och därmed vara bättre

öronvittnen. Här är dock forskningen inte enad. I några studier har blinda öronvittnen visat sig prestera bättre än seende öronvittnen (Granhag, 2001). Andra studier har visat att blinda öronvittnen inte presterar bättre än seende (Coob, Lawrence & Nelson, 1979; Winograd, Kerr & Spence, 1984).

Utöver de faktorer som Yarmey (1995) och Granhag (2001) nämner finns det naturligtvis andra faktorer som kan påverka ett vittne, exempelvis att få auditivt och visuellt stimuli samtidigt. Forskning visar att människor är vana att minnas ansiktet på folk som ett igenkänningsätt, detta underlättar förstås ögonvittnens uppgift. Att förlita sig till en röst ter sig inte lika självklart. Igenkänning enbart på röst görs således endast om visuella ledtrådar inte är möjligt (Cook & Wilding, 2007; Philippon et al., 2007). Studier har visat att det är en fördel att se ansikte och höra en röst samtidigt, när det gäller att känna igen ansikten. Det stjälpes emellertid att se ett ansikte om man är ute efter att få en pålitlig röstigenkänning (Cook & Wilding, 2007; Stevenage, Howland & Tippelt 2011). Dessa resultat tyder på att människan har en selektiv förmåga när det gäller störningar. Människor väljer att fokusera på det visuella istället för det auditiva. Stevenage et al. (2011) poängterar också att oavsiktlig respektive avsiktlig inläring av röst kan avgöra. Vet en person att denne ska lyssna till en röst kan personen förbereda sig, vet man det inte kan ansiktet tillåtas störa röstinläringen (Stevenage et al 2011). Cook och Wilding kallar detta ansiktsöverskuggning eller en generell visuell överskuggning (Cook & Wilding, 2007). Det visuella intrycket man får dominerar, i ett asymmetriskt mönster av störningar menar Stevenage et al. (2011). Yarmey (1986) fann dock att vittnen till en skådespelad våldtäkt varken var bättre eller sämre på att minnas röster oavsett om det var natt eller dag. Det fanns många fler detaljer som kunde störa uppmärksamheten för röstigenkänningen under ”dag-förhållandet”. Yarmey (1986) resultat går därmed emot Cook och Wilding (2007) samt Stevenage et al. (2011) slutsatser.

Konfidensbedömning

En viktig aspekt inom vittnesforskning är hur säker ett vittne känner sig på sitt utpekande, så kallad konfidens. Mycket av den tidigare forskning som gjorts avseende konfidens har visat på ett svagt eller icke existerande samband mellan hög konfidens hos vittnet och huruvida vittnets utpekande är korrekt eller ej enligt Yarmey (1995). Detta har även framkommit i andra studier, både avseende röstigenkänning och ansiktsigenkänning bland annat hos Olsson, Juslin och Winman (1998), Read och Craik (1995) och Yarmey (1986).

Konfidens som predicerar korrektheten har visat sig vara starkare vid ansiktsigenkänning än vid röstigenkänning (Olsson et al., 1998). Olsson et al. (1998) menar att deltagarna kunde vara realistiska när det gällde ansiktsigenkänningen, men för röstigenkänning var resultatet extremt dåligt och deltagarna visade en stor överkonfidens. Olsson et al. (1998) föreslår att skillnaden i predicering av röst och bild kan ligga i laboriemiljön i sig. Laboriemiljön är mycket annorlunda än annan, vanlig miljö och deltagarna kan ha vaggats in i en falsk trygghet att de bör kunna prestera samma i laboriemiljön som i vanliga livet (Olsson et al., 1998). Olsson et al. (1998) visade även på att användande av olika strategier kunde hjälpa deltagarna. Vissa strategier passade sig bättre för bilder respektive för röster.

Ett något annorlunda resultat fick Stevenage et al. (2011) i sin studie. Studien tittade på röst och ansiktsigenkänning men också med betingelsen att en röst kunde följas av en bild på ett ansikte eller tvärtom. Bäst lyckades deltagarna predicera sin korrekthet när de enbart såg ansikten och inte fick därpå följande röster. Stevenage et al. (2011) fann överlag i sin studie att självuppskattad konfidens var relaterad till korrektheten för ansiktsigenkänning. För röstigenkänning var denna korrelation emellertid svag eller icke existerande (Stevenage et al., 2011). När deltagarna presenterades för kortare röstklipp är det större korrelation mellan konfidens och korrekthet (Stevenage et al., 2011).

Forskning har också visat att amerikanska jurister tenderar att överestimera konfidens med korrekthet samt ha en tendens att tro att de som varit utsatta för brutala brott har bättre minnesförmåga, än de som blivit utsatta för mindre brutala brott (Philippon et al., 2007). Samma studie visar emellertid att generell kunskap om röstigenkänning är bättre än väntat. Deltagarna i kontrollgruppen uppvisar relativt låga nivåer av konfidens. Deltagarna visade genom sina svar att de var medvetet okunniga inom vissa områden (Philippon et al., 2007). Experter har emellertid inte visat sig vara bättre än gemene man att förstå sambandet mellan konfidens och korrekthet, synnerligen inte när det gäller våldsbrott, tidsuppfattning samt äldre ögonvittnen. Inte heller har poliser (experter) visat sig ha bättre kunskap om vittnens minne, som de själva tenderar att tro de har (Philippon et al., 2007).

Rättsväsendets syn på öronvittnen

Det går att prata mycket om minnets egenskaper och hur en försöksperson klarar sig i laboriemiljö. Men det är viktigt i en studie som denna att beskriva hur det praktiska arbetet med öronvittne ser ut så väl som det teoretiska. I en rapport från Rikspolisstyrelsen (2005) framkommer det att det i regel är svårare att minnas en röst än ett ansikte och att det i

Sverige inte finns någon utarbetad metod för hur en röstkonfrontation bör gå till.

Rikspolisstyrelsen rekommenderar emellertid att det iakttas samma principer som vid ögonvittnen, såsom att vara noga med att välja figuranter som speglar den eventuella gärningsmannen, i ålder, kön etc. Rikspolisstyrelsen skriver också att det är fullt i sin ordning ifall vittnen vill lyssna på rösterna flera gånger. Samt att alla som medverkar ska säga samma sak. Här framkommer dock en viss diskrepans från Granhag och Christianson (2008). De menar att det är en missuppfattning att alla personerna vid en röstkonfrontation bör läsa samma sak vilket blir onaturligt och problematiskt. Detta går emot vad Granhag (2001) beskrivit, då Granhag konstaterade att alla figuranter bör yttra samma fras, men inte nödvändigtvis samma fras som gärningsmannen. Detta motiverade Granhag (2001) med att användandet av samma fras som gärningsmannen sa vid tillfället är endast till hjälp om alla bidrag till röstkonfrontationer håller samma stressnivå och intonationsmönster som gärningsmannen hade. Detta följer av att om vittnet hör samma fras men i neutral ton kommer det att försvåra identifikationen. Detta motiverades även av att forskning visat att vittnen baserar sin identifikation på rösters grad av aggressivitet snarare än på övrig särprägel (Granhag, 2001).

Det framstår som ett problem att polisen, som är de som genomför förhören med öronvittnena i praktiken, har en uppfattning och att forskare har en annan uppfattning bland annat om vilken metod som är bäst att använda; om alla deltagare bör yttra samma fras eller inte.

Utifrån de resultat Yarmey (1995) presenterar framkommer det att de flesta försöksdeltagarna hade en föreställning om att de är mycket bättre på att minnas röster än vad de egentligen är. Studien som Yarmey (1995) tar upp visar därmed att det finns en övertro bland människor på sin förmåga att minnas röster. Det är farligt menar Yarmey eftersom, det i alla fall från ett amerikanskt perspektiv, finns en möjlighet att dessa personer kan bli valda att sitta i en eventuell jury. Risken blir att det sitter personer i en eventuell jury som har en övertro på ett vittnes förmåga att minnas en röst, exempelvis som i Lindbergh-fallet som presenterades i början. Det kan också kopplas till problemet med konfidensbedömning. Philippon et al. (2007) fann dock ett något annorlunda resultat i sin studie. Studien undersökte vilken kunskap engelska poliser och lekmän besitter när det kommer till öronvittnen. Bland annat valde de att titta på konfidensbedömning och vilken tilltro som gick att sätta till om vittnen sa att det var helt säkert på att den pekade ut rätt röst, samt om försöksdeltagarna förstod hur svårt det faktiskt var att minnas röster. Studien av Philippon et al. (2007) visade till skillnad från Yarmey (1995) att försöksdeltagarna faktiskt hade en ganska god kunskap om

svårigheten att peka ut en gärningsman utifrån rösten samt att konfidens inte är samma sak som korrekthet. Det som dock är oroväckande var att poliserna ibland hade sämre kunskap än lekmännen, framförallt verkade poliserna i studien sätta större tro till ett vittnes konfidens angående ett korrekt utpekande (Philippon et al., 2007).

Expertvittne

Möjligheten att det skulle finnas vissa personer eller yrkesgrupper som är bättre på att minnas detaljer vid överfall, och därmed vara bättre vittnen är omdebatterat. Olika studier har visat på olika saker, både när det gäller ögonvittnen och öronvittnen. Christianson, Karlsson och Persson (1998) genomförde en studie där två grupper av poliser testades, en grupp med väldigt erfarna poliser, en grupp med polisaspiranter samt en grupp högstadielärare och en grupp psykologistuderenter. Grupperna fick se ett simulerat brott via ett antal bilder som visades för dem, därefter fick de vittna om vad de sett för något. Överlag så visade det sig att de yrkeserfarna poliserna var något bättre än de övriga grupperna exempelvis genom att komma ihåg fler antal perifera detaljer. På vissa delar av vittnesmålet var emellertid psykologistudenterna lika bra som de yrkeserfarna poliserna, bland annat när gärningsmannens utseende skulle beskrivas. Christiansson et al. (1998) argumenterar för att det är de erfarna polisernas yrkeserfarenhet som gör att de är något bättre på att minnas saker jämfört med de övriga grupperna.

Vidare så presenterar Borgström (2007) i sin studie ett resultat som visar på att den expertgruppen som hon valde att titta på; poliser, väktare, kriminalvårdare och larmoperatörer, inte var bättre än kontrollgruppen utan rent av uppvisade sämre resultat i några aspekter. Borgströms (2007) tanke med sin undersökning var att se om den expertgruppen hon hade skulle ha bättre minne för ansikte och röster jämfört med kontrollgruppen. Denna tanke baserade hon på att de genom sin yrkeserfarenhet förvärvade kunskap som eventuellt skulle kunna göra dem till bättre vittnen än den kontrollgrupp hon använde sig av, då kontrollgruppen inte alls hade någon form av yrkeserfarenhet när det kom till identifikationsuppgifter. Borgström (2007) använde sig av två test, dels ögonvittne dels öronvittne. Båda testen följde samma mönster, exponering för figuranten följt av en distraktionsuppgift och avslutades med en lineup där försöksdeltagaren skulle peka ut figuranten. Som nämndes ovan var experimentgruppen sämre än kontrollgruppen på vissa aspekter såsom ansiktsigenkänningen vilket går emot det resultat som Christianson et al. (1998) hittade i sin studie. Dock fann Borgström (2007) att hennes experimentgrupp var signifikant bättre på att identifiera röster. Detta är ett intressant resultat för det har länge

debatteras inom forskningen om minne för röster, om det finns grupper som kan vara bättre än andra på att minnas röster. Det som gör resultatet intressant är att tidigare forskning har haft svårt att hitta grupper som är bättre än andra grupper på att minnas röster, bland annat avfärdar Hollien (2002) påståendet att personer som är tränad i fonetik skulle vara bättre på att minnas röster. Även Clifford (1980) menar att det verkar vara svårt att hitta grupper som skulle vara bättre på att minnas röster än andra grupper. Borgströms (2007) resultat visade dock på att det faktiskt kan vara så att det finns grupper, eller yrken, som kan ha lättare än andra grupper att minnas röster. Borgström (2007) resultat visar även att expertgruppen var mer underkonfident än kontrollgruppen. I sin diskussion efterlyste Borgström (2007) mer forskning i området, gärna med personer som satt mycket i telefon exempelvis kundtjänsttelefonister och larmoperatörer.

Syfte och hypoteser

I ljuset av den tidigare forskning som presenterats här så framstår den kunskap som forskarvärlden besitter som väldigt snårig. Det ena forskningsresultatet pekar åt ett håll medan ett annat pekar lite åt ett annat håll. I det här fallet kan den osäkerhet som finns i ämnet få förödande konsekvenser. Som nämnts tidigare finnas det inte alls så mycket forskning kring öronvittnen som det finns om ögonvittnen, och det märks bland annat i den diskrepans som finns mellan den svenska rikspolisstyrelsens riktlinjer för hur en vittneskonfrontation bör gå till och vad Granhag och Christianson (2008) rekommenderar. Det är också viktigt att uppmärksamma det glapp som finns mellan hur säkert ett vittne säger sig vara på att de gjort ett korrekt utpekande och om utpekandet verkligen stämmer, så kallad konfidensbedömning. Som ovan presenterat finns det väldigt liten korrelation mellan hög konfidens och korrekthet i utpekandet, både för öronvittnen och för ögonvittnen. Därför måste poliser och övriga delar av rättsapparaten vara noga med att inte lägga allt för stor vikt vid ett sådant uttalande om hög konfidens.

Syftet med denna studie är precis som nämndes inledningsvis att försöka bidra till den kunskapsbas som finns om öronvittnen. Studien kommer att fortsätta i samma spår som Borgström (2007) att undersöka om det finns vissa yrkesgrupper som är bättre än andra på att minnas röster och därmed fungera som bättre vittnen. Eftersom det inte finns så mycket forskning som pekar på att vissa människor, eller yrkesgrupper är bättre på att göra korrekta utpekanden när det handlar om öronvittnen är det extra intressant att fortsätta på samma linje som Borgström (2007), för att ytterligare utreda denna företeelse. Borgström (2007) hade i sin undersökning valt ut poliser, väktare och larmoperatörer som sin expertgrupp eftersom de

genom sitt yrke kom i kontakt med fler identifikationsuppgifter än vad gemene man gör. I denna studie kommer enbart larmoperatörer utgöra expertgruppen. Tanken med detta är att larmoperatörer i sitt arbete kommer i kontakt med ett så stort antal röster under ett arbetspass att de utvecklar en finkänslighet som underlättar för dem att minnas eventuella röster de kommit i kontakt med. Även det bitvis höga tempot i deras arbetsmiljö, som kräver att de måste hålla många saker i huvudet samtidigt kan tänkas bidra till deras minne. I Borgströms (2007) studie visade det sig att expertgruppen var signifikant underkonfident jämfört med den kontrollgrupp som användes, detta är något som vi tror kommer gälla för vår expertgrupp i denna studie också. Underkonfidensen kommer att ligga i att expertgruppen eftersom de utsätts för så många samtal under sitt arbete kommer att inse att det är en svårare uppgift att minnas än vad kontrollgruppen kommer att göra. Vilket kommer leda till att expertgruppen kommer ha en total konfidens som är lägre än kontrollgruppen oberoende av om deras utpekande är korrekt eller inte. I studien kommer försöksdeltagarna även att testas på ansiktsigenkänning, detta för att kunna veta att det är just röster som expertgruppen är bättre på och inte att de har ett bättre minne generellt. Slutligen kommer även fenomenet med verbal överskuggning att testas, tanken är att se om expertgruppen genom sin yrkeserfarenhet är skyddad från detta fenomen till skillnad från kontrollgruppen, när det gäller att minnas röster.

Studien ska ge svar på följande hypoteser:

H_1 = Expertgruppens yrkeserfarenhet ger en bättre förmåga till att kunna diskriminera mellan gamla och nya röster.

H_2 = Expertgruppen kommer vara mindre konfident i sina utpekningar för röster.

H_3 = Expertgruppens yrkeserfarenhet kommer att minska risken för verbal överskuggning för röster.

Metod

Undersökningsdeltagare

Experter: Expertgruppen rekryterades från en larmcentral tillhörande ett av världens största bevakningsföretag. Expertgruppen bestod av 24 stycken yrkesverksamma personer varav 7 kvinnor och 17 män, i åldrarna 20-60 år ($M=31.6$ år). Expertgruppen rekryterades utifrån tanken om att det arbetet de gör främjar möjligheterna för att utveckla bättre förmåga att minnas röster. På larmcentralen i fråga får varje operatör ta emot eller ringa ca 200 samtal per arbetspass, ofta under ett väldigt högt arbetstempo. Utifrån detta kriterium valdes försöksdeltagarna ut till expertgruppen. Kontrollgruppen valdes ut genom bekvämlighetsurval och bestod av 24 personer, 14 kvinnor och 10 män, i åldrarna 20-65 år ($M=34.45$ år). Ingen

av dessa personer var yrkesverksamma på den larmcentral där expertgruppen rekryterades eller hade någon känd yrkeserfarenhet som kunde jämföra dem med expertgruppen. Samtliga försöksdeltagarna hade svenska som modersmål. Ingen av försöksdeltagarna uppbar ersättning för sitt deltagande.

Material

Ansiktsbilder. I den delen av experimentet som testade ansikte användes 96 bilder på olika ansikten, varav hälften var kvinnor och hälften män. 16 stycken av bilderna användes i den inledande övningsdelen och 80 stycken av bilderna användes sedan i experimentets olika block. Samtliga figuranter på bilderna som användes var fotograferade med både ansikte och kropp vända mot kameran, både figuranternas ansikte och axlar syntes. Samtliga bilder var tagna mot en vit bakgrund. Bilderna bestod av människor med varierande etniskt ursprung, hudfärg och ålder. Ingen av försöksdeltagarna hade någon känd kontakt med figuranterna som användes i ansiktsexperimentet.

Ljudklipp. De ljudklipp som användes i studien var samtliga hämtade från en studie av Borgström (2007). Ljudklippen var inspelade med en MP3-spelare av märket T.Sonic TM 630. För att underlätta inspelningen och ljudkvaliteten användes en mindre extern mikrofon, en så kallad slipsmikrofon av märket MICTC2. 52 ljudklipp användes i testningen, av de ljudklippen bestod figuranterna av 26 kvinnor i åldrarna 16-65 år medan 26 av ljudklippen bestod av män där åldern varierade mellan 20-65 år. 8 stycken av ljudklippen användes i den inledande övningsdelen och de övriga 44 användes i resten av experimentet. Samtliga ljudklipp bestod av personer som hade svenska som modersmål och hade dialekter som är hemmahörande i södra Sverige, främst Kristianstadtrakten. Samtliga figuranter läste samma text bestående av 41 ord. Texten lød:

Människor är i grunden inte särskilt olika, men i ett krisläge finns det dels de som blir hysteriska och bara gråter, dels de som omedelbart börjar agera för att försöka lösa problemen. Det gemensamma för båda sorterna är att återfå kontrollen (Werkelin & Svanberg, 2006, sid. 128).

Figuranterna läste texten i normal samtalston och inläsningshastigheten varierade något mellan figuranterna. För att skapa så bra klipp som möjligt redigerades samtliga ljudklipp i ljudredigeringsprogrammet Audacity. Ingen av försöksdeltagarna hade någon känd kontakt med figuranterna som användes i röstexperimentet.

Procedur

Samtliga försöksdeltagare genomförde testet under liknande premisser. Antingen utfördes testet i försöksdeltagarnas hemmiljö/arbetsplats eller på ett grupprum i Hus P (Psykologen) på Lunds Universitet. Samtliga platser där försöksdeltagarna testades var lugn och avskild för att minska eventuella störande moment. Försöksdeltagarna hade inte någon tidsgräns på sig när de genomförde testet. Experimentet tog i genomsnitt ca 50 min att genomföra.

Experimentet konstruerades i mjukvaruprogrammet E-prime. Försöksdeltagarna genomförde experimentet sittandes framför en bärbar dator där E-prime fanns installerat. På skärmen blev försöksdeltagarna presenterade för experimentet via skriftliga instruktioner.

Försöksdeltagarna fick läsa igenom instruktionerna i egen takt och blev tillfrågade ifall de hade några frågor innan de olika delarna av experimentet inleddes. Försöksdeltagarna lämnade sina svar genom att trycka på färgmarkerade tangenter på datorns tangentbord.

Experimentet bestod av fyra stycken block. Samtliga försöksdeltagare utförde samtliga testblock. Två stycken av blocken var så kallade ansiktsblock, det vill säga block som var avsedda att testa försöksdeltagarens minne för ansikte. De två andra blocken var så kallade röstblock, vilket betydde att de var avsedda att testa försöksdeltagarnas minne för röster.

Varje del av experimentet, röst och ansiktsdel, bestod av en inledande övningsdel som användes för att säkerställa att försöksdeltagarna förstod uppgiften. Varje block i experimentet, bestod av en instuderingsfas, en mellanliggande uppgift samt en testfas.

Uppgiften i instuderingsfasen var att bedöma om rösten eller ansiktet som presenterades tillhörde en man eller en kvinna. Bilden presenterades i två sekunder och rösten pratade under femton sekunder. Efter instuderingsfasen fick deltagaren en av två uppgifter, antingen fick de till uppgift att beskriva den senaste rösten eller det senaste ansiktet de blivit presenterade för.

Försöksdeltagarna bads att skriva ner sin beskrivning så detaljerat som möjligt på ett skrivhäfte de hade tillgång till. Denna uppgift varade under 5 min och var tänkt att testa försöksdeltagarnas minne samt eventuella effekter av verbal överskuggning. Den andra uppgiften som försöksdeltagarna kunde bli ombudda att utföra var att räkna upp så många svenska kommuner som möjligt under 5 min och skriva ner dessa i ett skrivhäfte. Den andra mellanliggande uppgiften var tänkt att dels testa försöksdeltagarnas minne samt att fungera som kontroll för effekterna av verbal överskuggning. Efter att den mellanliggande uppgiften var avslutad inleddes testdelen.

Uppgiften i testdelen var att beskriva huruvida det ansikte eller röst försöksdeltagarna blev presenterad för hade förekommit i instuderingsdelen, med andra ord om ansiktet eller rösten var nytt eller gammalt. Testdelen inleddes alltid med ett antal så kallade fillerröster eller

filleransikte, stimuli som inte varit med i instuderingsfasen, 3 stycken i ansiktsdelen och 2 stycken i röstdelen. Precis som under instuderingsfasen presenterades bara ansiktet under två sekunder och rösten under femton sekunder. Varje bedömning följdes av en konfidensbedömning. Konfidensbedömningen var indelad i tre steg, osäker, ganska säker och helt säker. De försöksdeltagare som hade fått till uppgift att inledningsvis göra det så kallade ”verbaliseringsblocket”, det block där de fick till uppgift att beskriva det ansiktet eller den rösten de hörde, fick därefter genomgå ”kontrollblocket”, det vill säga det blocket där de blev ombedda att räkna upp så många svenska kommuner de kunde komma på, och tvärtom. Blocken alternerades mellan försöksdeltagarna så att hälften fick börja med ”verbaliseringsblocket och hälften fick börja med ”kontrollblocket”. Även ordningen på röstdelen och ansiktsdelen alternerades för att motverka eventuella övningseffekter.

Resultat

Totalt utförde de 48 försöksdeltagarna 3840 identifikationsuppgifter för ansiktsigenkänning och 1920 röstigenkänningsuppgifter. Medelvärde samt standardavvikelsen har räknats om till en proportion av antalet korrekt respons, pga. olika antal frågor i experimentet, 40 frågor totalt för varje försöksdeltagare för ansikten respektive 20 frågor totalt för röster. Proportionen har räknats ut genom att subtrahera felaktiga utpekningar (false alarm) från korrekta utpekningar(hit), därmed fås ett värde på deras minnesprestation fram, ett så kallat PR-värde, från 0-1,0. De inledande filler- ansiktena och rösterna, 3 ansikte och 2 röster har inte använts för att kalkylera försöksdeltagarnas minnesprestation. Som framgår i *tabell 1* hade båda grupperna en högre minnesprestation i de båda ansiktsblocken än i röstblocken. Expertgruppen presterar något bättre än kontrollgruppen i verbaliseringsblocket för röster, 0,2 jämfört med 0,13. Dock är inte expertgruppen bättre i icke-verbaliseringsblocket för röster, 0,13 jämfört med 0,16.

Tabell 1. Deskriptiv statistik för genomsnittlig minnesprestation för samtliga block.

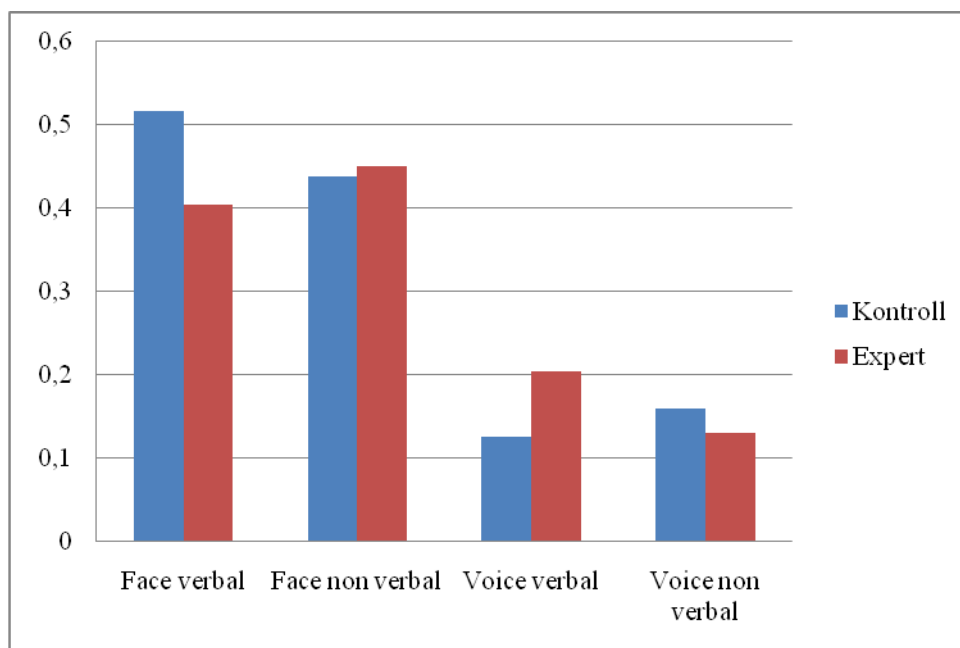
Proportion av PR värde

	N	Face Verbal		Face Non-Verbal		Voice Verbal		Voice Non-Verbal	
		M	SD	M	SD	M	SD	M	SD
Expert	24	0,40	0,23	0,45	0,26	0,20	0,29	0,13	0,27
Kontroll	24	0,52	0,27	0,44	0,33	0,13	0,27	0,16	0,28

En mixad tvåvägs-ANOVA genomfördes för att undersöka om det fanns några signifikanta skillnader mellan expertgruppen och kontrollgruppen avseende två olika faktorer.

Den första av faktorerna var Domän, som avser skillnaden i förmågan att minnas röster vs ansikte. Den andra faktorn som användes var Verbalisering, som avser förmågan att minnas under verbaliseringsblocken vs ickeverbaliseringsblocken. Kravet för varianshomogenitet var uppfyllt för samtliga variabler enligt Levene's Test of Equality of Error Variances.

Det fanns ingen signifikant interaktionseffekt mellan Expert och Domän, $F(1, 46) = 1.128, p = .29$. Vilket talar för att expertis inte gav en förhöjd minnes prestation. Det fanns heller ingen signifikant interaktionseffekt mellan Expert och Verbalisering, $F(1, 46) = .01, p = .92$. Ingen signifikant interaktions effekt fanns heller mellan Domän och Verbalisering, $F(1, 46) = .005, p = .95$. Vilket pekar på att det troligtvis inte har någon effekt huruvida verbalisering förekommer oavsett domän. Inte heller fanns det någon statistisk signifikant interaktionseffekt mellan Expert, Domän och Verbalisering, $F(1, 46) = 3.53, p = .07$. Dock fanns det en statistik huvudeffekt avseende Domän, $F(1, 46) = 71.17, p = <.0005$. Vilket talar för att försöksdeltagarna hade lättare för att känna igen ansikte än röster, för en tydligare överblick på huvudeffekten se *figur 1*. Det fanns inte någon signifikant huvudeffekt för Expert, $F(1, 46) = .07, p = .8$ eller för Verbalisering, $F(1, 46) = .196, p = .66$.



Figur 1. Genomsnittlig minnesprestation presenterad i proportion för samtliga block.

För att utreda den trenden $F(1, 46) = 3.53, p = .07$ som syns i trevägsinteraktionen mellan expert, domän och verbalisering, utfördes ytterligare test. Baserat på *figur 1* kan det framförallt vara intressant att titta närmare på eventuella skillnader mellan verbalisering och icke-verbalisering på de olika betingelserna. Resultaten av de paired t-tests som genomfördes

visade att det inte fanns någon signifikant skillnad, varken för kontrollgruppen eller för expertgruppen, mellan verbaliseringen och icke-verbaliseringen. Detta gällde på både röstblocken och ansiktsblocken. Se *tabell 2*.

Tabell 2. Resultat av paired t-tests för skillnaden mellan verbalisering och icke-verbalisering på samtliga block.

	Face Verbal- Face Non Verbal		Voice Verbal- Voice Non-Verbal	
	T	p	T	P
Expert	-.69	.49	1.1	.27
Kontroll	.92	.37	-.43	.67

Konfidens

Testet bestod även av 43 konfidensuppgifter på varje block på ansiktsdelen, verbal samt icke verbal, varav de tre första var fillers. Testet bestod av 22 konfidensuppgifter på varje block på röstdelen, verbal samt icke verbal, varav de två första var fillers. Försöksdeltagarna hade tre valmöjligheter för att uppskatta sin konfidens; osäker som gav 0 poäng, ganska säker 1 poäng samt helt säker som gav 2 poäng. Försöksdeltagarnas totala konfidens räknades ihop, separat för varje block. Sedan gjordes en uträkning på hur varje försökspersons totala konfidens såg ut i proportion till antalet uppgifter på respektive block, där 1,00 var helt säker på alla utpekningar och 0,00 var osäker på alla utpekningar. Detta proportionerliga värde fick fungera som mått för varje försöksdeltagares totala konfidens på de olika blocken. Det visade sig att båda grupperna hade totalt sett en högre konfidens avseende ansiktsblocken än röstblocken. Samt att expertgruppen hade en genomsnittligt sett högre konfidens än kontrollgruppen avseende röster, 0,65 och 0,63 jämfört med 0,55 och 0,52 för kontrollgruppen, se *tabell 3*.

Tabell 3. Deskriptiv statistik för genomsnittlig total konfidens presenterad för samtliga block.

Proportion av PR värde

	N	Face Verbal		Face Non-Verbal		Voice Verbal		Voice Non-Verbal	
		M	SD	M	SD	M	SD	M	SD
Expert	24	0,76	0,16	0,79	0,13	0,65	0,17	0,63	0,18
Kontroll	24	0,77	0,15	0,79	0,14	0,55	0,19	0,52	0,21

En mixad trevägs-ANOVA användes för att utreda skillnaderna i konfidensbedömningarna mellan grupperna. Huruvida försöksdeltagaren tillhörde kontroll

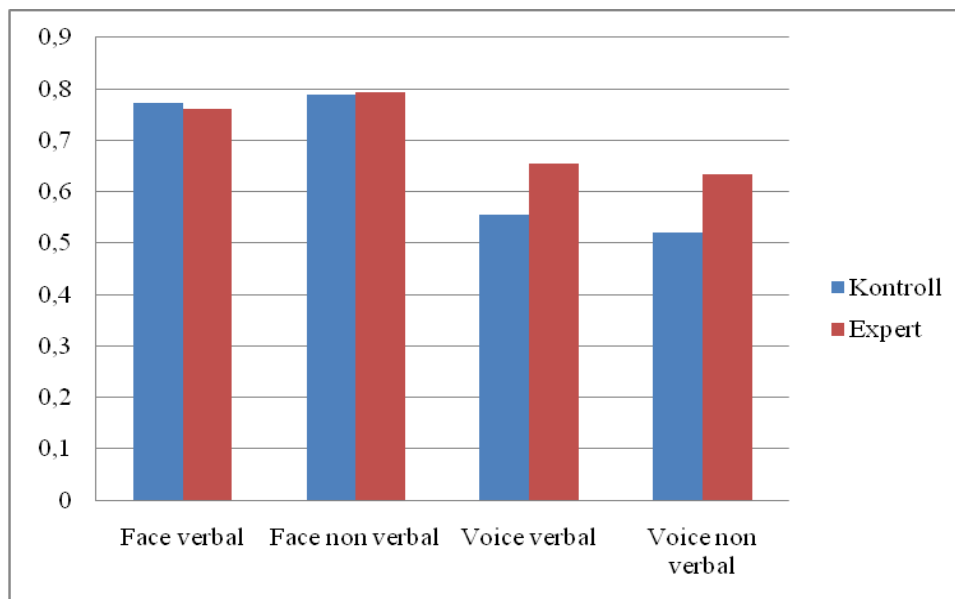
eller expertgruppen var en mellangrupsvariabel medan Domän (förmågan att minnas röster eller ansikte) och Verbalisering (förmågan att minnas under verbaliseringsblocken vs ickeverbaliseringsblocken) var inomgruppsvariabler. Resultatet från den mixade trevägs-ANOVA presenteras i *tabell 4*.

Tabell 4. *Trevägs-ANOVA för konfidensbedömning gällande Expert, Domän och Verbalisering*

Faktor	F	P
Domän	111,093	<0.0005*
Domän * Expert	9,327	0.004*
Verbalisering	0.20	0.887
Verbalisering * Expert	0.433	0.514
Domän * Verbalisering	6.886	0.012*
Domän*Verbalisering*Expert	0.25	0.874
Expert	1.481	0.23

* $p < 0.05$

Som framgår i *tabell 4* finns det en rad intressanta fynd i trevägs-ANOVA:n. Det fanns en signifikant interaktionseffekt mellan Domän*Expert $F(1,46) = 9,327$, $P = 0.004$. Denna interaktionseffekt pekar på att expertgruppen hade större total konfidens vid utpekande av röster än vad kontrollgruppen hade men inte när det gällde ansikte. Interaktionseffekten tydliggörs i *figur 2*. Det fanns även en signifikant interaktionseffekt mellan Domän*Verbalisering $F(1,46)$, $P = 0.012$. Interaktionseffekten mellan Domän*Verbalisering beror på att försöksdeltagarna hade ett högre medelvärde på ansiktsblocket som inte innehöll verbalisering medan det var precis tvärtom för röstblocken, se *tabell 4*. Även en signifikant huvudeffekt stod att finna, denna gällde Domän $F(1,46)$, $P = <0.0005$. Huvudeffekten klagörs i *figur 2*, där både expert och kontrollgrupp hade lägre konfidens på röstblocken, oavsett om verbalisering förelåg eller ej.



Figur 2. Genomsnittlig total konfidens presenterad i proportion för samtliga block.

Diskussion

Syftet med denna studie var att se om en viss typ av yrkeskunskap bidrog till en bättre förmåga att göra korrekta igenkänningar av röster. Expertgruppen som användes i studien förväntades prestera bättre i förhållande till kontrollgruppen eftersom deras yrke gjort att de skapat sig erfarenheter som bidrog till att främja minnet för röster, exempelvis främja de olika minnesprocesserna inkodning, lagring och framplöckning. Däremot förväntades kontrollgruppen och expertgruppen inte skilja sig åt avseende förmågan att minnas ansikte. Expertgruppen förväntades också vara mindre konfident i sina utpekningar eftersom deras yrkeskunskap skapade en större respekt för svårigheten att göra korrekta utpekande baserat på röst. Slutligen förväntades expertgruppen vara mer skyddad i den betingelsen som undersökte hur verbal överskuggning påverkade minnet för röster.

I studien visade det sig att experterna inte gjorde bättre ifrån sig på de block av experimentet som testade minnet för röster. Det går emot den inledande hypotesen, att experter skulle vara bättre än kontrollgruppen. Detta resultat kan vara ett argument mot att det arbete som experterna i denna grupp utför förbättrar deras minne för röster. Att expertgruppens yrke inte bidrog till att främja deras episodiska minne av rösterna eller underlättade för deras minnesprocesser jämfört med kontrollgruppen. Detta resultat är i linje med vad Hollien (2002) och Clifford (1980) diskuterade avseende expertis för röstminne. Resultatet i studien går dock emot de resultat som Borgström (2007) fick. I Borgströms (2007) studie visade expertgruppen en signifikant bättre förmåga att göra korrekta utpekande av röster. Det ska dock tilläggas att Borgström (2007) hade ett annat upplägg på sin studie

samt en annan typ av expertgrupp. Försöksdeltagarna hos Borgström (2007) fick göra sitt utpekande utifrån en line-up istället för som i denna studie att bedöma var röst för sig. I Borgström (2007) bestod expertgruppen inte bara av larmoperatörer utan till stor del av väktare och poliser, med argumentet att de hade vana av identifikationsuppgifter. Måhända är detta faktorer som kan förklara skillnaden i utfallet mellan de två olika studierna.

Resultatet visar ingen signifikant skillnad mellan grupperna i de block som innefattade ansikte. Detta betyder att kontrollgruppen och expertgruppen hade en likvärdig förmåga att minnas ansikte de sett. Vilket får ses som ganska plausibelt eftersom det inte finns så mycket som talar för att den yrkeserfarenhet som larmoperatörer förvärvat skulle bidra till att öka deras minnesprestation för ansikte.

Resultatet avseende konfidens föll inte heller ut enligt den hypotes som var uppställd. I studien låg fokus på försöksdeltagarnas totala konfidens, hur säkra de kände sig på sina utpekningar. Enligt den andra hypotesen i studie skulle expertgruppen, med hjälp av den erfarenhet som de förvärvat, totalt sett ha en lägre konfidens på röstblocken. Tanken bakom hypotesen var att yrkeserfarenhet skapade en större respekt för uppgiften att korrekt kunna bedöma huruvida rösten som presenterades är ny eller gammal. Experimentets resultat visade på att det fanns en signifikant interaktionseffekt mellan domän och grupp avseende konfidens. Interaktionseffekter visade tvärt emot hypotesen att expertgruppen var mer konfident än kontrollgruppen i sina utpekningar för röster. När det gällde utpekningar för ansikte fanns det ingen signifikant skillnad. Resultatet pekade på att expertgruppen istället för att vara mer försiktiga istället blev mer säkra i sina utpekanden. Måhända trodde de att deras yrkeserfarenhet gjorde att de skulle klara uppgiften bättre. Eftersom expertgruppen inte presterade bättre än kontrollgruppen talar dock detta för att expertgruppen var överkonfident i sina utpekningar. Resultatet går i linje med Yarmey (1995) som pekade på att folk lätt överskattar sin förmåga att göra korrekta utpekande. Emellertid går resultatet emot Borgström (2007), som kom fram till att kontrollgruppen var mer konfident än expertgruppen. Återigen är det dock viktigt att påpeka att Borgström (2007) använde sig av en annan typ av expertgrupp, där bland annat ett antal poliser och väktare ingick. Kanske var expertgruppen i Borgström (2007) mer försiktig i sina uttalanden, eftersom de i egenskap av sina jobb inser vilka följder ett felaktigt utpekande av en gärningsman kan få i verkligheten. Dock fann Philippon et al. (2007) att poliserna i deras studie satte större tilltro till vittnens konfidens än vad lekmän gjorde. Det kan eventuellt bero på att det är skiljda saker att bedöma värdet av ett vittnes konfidens och en annan sak att själv göra utpekandet.

Vad gäller den tredje hypotesen, att expertgruppens yrkeserfarenhet skulle minska risken för verbal överskuggning för röster visade resultatet inte heller några signifikanta skillnader. Experter var därmed inte mer skyddade för verbal överskuggning jämfört med kontrollgruppen vare sig för ansikten eller för röster. Detta kan eventuellt förklaras med att, testet ansågs vara svårt överlag eller att individuella skillnader förelåg. Att beskriva ett ansikte eller en röst är inte något som är vanligt i ens vardagsliv, därför hade man kunnat tänka sig att enligt innehållsteorin (Chin & Schooler, 2008) skulle testdeltagarna drabbas av verbal överskuggning. Testdeltagarna hade några minuters betänketid innan det automatiserade datatestet gick vidare efter verbaliseringen. Eventuellt upplevde deltagarna detta som en forcerad återkallelse. Detta leder ju till att mer felaktig information kommer fram och verbal överskuggning drabbar i högre grad (Chin & Schooler, 2008). Varför verbal överskuggning ändå inte skedde kan bero på att testdeltagarna hade lägre perceptuella kunskaper än verbal förmåga och därmed störde inte verbaliseringen deras minne av ansiktet eller rösten (Melcher & Schooler, 1996). En annan förklaring kan vara att kodningen och testet är likartat, som Vangas et al. (2005) fann kan verbal överskuggning minska om kodningen och testvillkoren är lika. Att det är samma fras som används genomgående i testet, som i denna studie, skulle därmed kunna minska risk för verbal överskuggning.

Utöver de hypoteser som ställdes upp framkom det även andra fynd som kan vara av intresse att föra fram och diskutera. Det fanns en signifikant huvudeffekt avseende minnesprestationen för domän. Huvudeffekten visade på att oberoende av grupp hade försöksdeltagarna svårare att minnas röster jämfört med ansikte. Denna kunskap kan vara av värdefull ur en rad olika perspektiv, både juridiskt och för framtida forskning. Eftersom den visar på att människor verkar ha svårare att minnas röster än ansikte. Det är intressant att notera att även på konfidensbedömningen fanns denna huvudeffekt avseende domän. Oavsett grupp hade försöksdeltagarna lägre konfidens på de block som berörde röst än de block som berörde ansikte. Vilket kan ses som att försöksdeltagarna själva upplevde röstblocken som svårare än ansiktsblocken. Eftersom hälften av gruppen gjorde ansiktsblocken först och andra hälften röstblocken först minimerades risken för eventuella tränings- och utmattningseffekter.

Den signifikanta interaktionseffekt som fanns mellan verbalisering och domän i resultatet beror på att försöksdeltagarna angav högre konfidens i ansiktsblocket som inte innefattade verbalisering jämfört med det ansiktsblock som innehöll verbalisering. Försöksdeltagarnas konfidens var precis tvärtom för röstblocken. Detta resultat kan tyda på att försöksdeltagarna kände sig mer säkra i sin utpekande när de verbaliserade för röster men inte när de verbaliserade ansikte. Kanske gav känslan av att beskriva en av rösterna de hört

intrycket av att uppgiften inte var lika svår som när de inte skulle verbalisera en av rösterna. Som presenterades i resultatet hade grupptillhörighet inte någon påverkan på denna interaktionseffekt, vilket framkom i den trevägsinteraktion som testades.

Den trend som uppmärksammades vid trevägsinteraktionen mellan grupp, domän och verbalisering i minnesprestation baserar sig utifrån *figur 1* på att försöksdeltagarna i de olika grupperna presterar bättre under olika betingelser i de olika blocken. Exempelvis har expertgruppen ett högre medelvärde på verbaliseringsdelen av röst men ett sämre på icke-verbaliseringsdelen av röstblocket, och tvärtom när det gäller ansiktsblocken. Det kan därför vara svårt att uttala sig säkert varför denna trend uppstått. De tester som utfördes gav inget signifikant resultat, när skillnaden mellan verbalisering och icke-verbalisering testades i paired t-tests för de båda grupperna. Det kan inte uteslutas att eventuella höga eller låga värden hos några av försöksdeltagarna kan ha påverkat medelvärdena eftersom studien hade ett begränsat antal försökspersoner.

De generella slutsatser som går att dra utifrån denna studie är framförallt att expertgruppen inte hade lättare för att minnas röster än vad kontrollgruppen hade. Expertgruppens yrkeserfarenhet verkade inte ge bättre *minnesprocesser*. Detta behöver emellertid inte betyda att det inte finns människor som har en bättre förmåga att minnas röster än vad gemene man har. Detta ska framförallt ses i ljuset av studien av Rämä och Courtney (2004) som visade på att olika delarna av hjärnan används beroende på om det är ansikte eller röster som försöksdeltagaren ska försöka minnas. Vilket kan göra att det finns människor som är bra på den ena uppgiften men inte på den andra. Ytterligare en slutsats som är av mer praktiskt betydelse för framförallt rättsväsendet är utfallet för konfidens. Både Philippon et al. (2007) och Yarmey (1995) varnade för att sätta allt för stor tilltro till ett vittnes konfidens och hur korrekt utpekande vittnet gjorde. Detta är en slutsats som också kan dras från denna studie, då framförallt expertgruppen angav en högre konfidens avseende röster men inte presterade signifikant bättre än kontrollgruppen. Enligt denna studie är det också så att det är svårare att peka ut röster jämfört med ansikte, något som också speglar sig i hur säker man är på sitt utpekande. Förhoppningsvis kan dessa slutsatser bidra till att utöka den kunskapsbas som finns om öronvittne och människors förmåga att minnas röster. För precis som Yarmey (1995) påpekar är det ett ganska litet område jämfört med exempelvis ögonvittnen. Studien försökte ge svar på ett ganska brett fält med frågor, vilket får ses som en styrka.

Att resultatet i uppsatsen inte föll ut enligt de hypoteser som ställts upp kan bero på en rad saker. Eftersom studien bestod av ett ganska litet antal försökspersoner kan detta ha påverkat. Eventuella skillnader mellan grupperna kan som nämnts tidigare ha försvunnit i eventuella

höga eller låga poäng från ett fåtal personer. Det kan vara så att det finns en skillnad mellan grupperna men att vårt urval var för litet för att hitta skillnader i minnesprestationen. En fördel med studiens expertgrupp var att till skillnad från exempelvis Borgströms (2007) var expertgruppen mycket mer homogen i förhållande till vilka arbetsuppgifter de utförde. Då Borgström (2007) använde sig av både poliser, fängvaktare, väktare och larmoperatörer.

Även om de båda försöksledarna var noga med att prata ihop sig innan experimenten utfördes, för att minska riskerna för testledarbias, går det inte att helt utesluta eventuella effekter. Dock var två försöksledare en förutsättning för att kunna nå ut till den expertgrupp som användes.

Precis som i alla studier av den här typen är det viktigt att poängtera att det är stor skillnad mellan ett experiment och verkliga livet. Det går inte att jämföra en experimentsituation med exempelvis något av de scenariorna som beskrivs i inledningen. Detta för att det inte är etiskt försvarbart att utsätta försöksdeltagarna för den stress och det obehag som är att förknippa med ett riktigt brott. Vilket gör att det inte går att kontrollera för vissa saker som kan påverka de olika minnesprocesserna, exempelvis den vapenfokuseringseffekt som Granhag och Christianson (2008) tar upp, som kan påverka inkodningen av ett minne på ett effektivt sätt. Detta är ett problem som alla studier som vill undersöka ögonvittnen eller öronvittnen dras med och som är viktigt att ha i åtanke när eventuella slutsatser ska dras. För framtida studier kan det dock vara av intresse att spinna vidare på den grund som lagts utifrån både Borgström (2007) och denna studie. Framförallt skulle ett större antal försöksdeltagare vara att föredra. Det skulle vara intressant att se resultatet av en studie som bestod av tre grupper, en kontrollgrupp, en grupp med larmoperatörer eller liknande samt en grupp med poliser. Detta för att ytterligare utreda den skillnad mellan expertgrupperna som uppstod i resultatet mellan Borgström (2007) och denna studie. Likadant med konfidensbedömningen. I en mer djupgående studie avseende konfidens hade det även varit intressant att göra en mer utredande undersökning för att försöka svara på varför resultatet föll ut som det gjorde. Det kan också vara intressant att titta på om resultaten hade blivit olika om andra typer av röster hade använts, eftersom denna studie nästan uteslutande använde sig av röster från nordöstra Skåne. Det hade även varit intressant att i en studie som denna med en expertgrupp och en kontrollgrupp, även testa en tredje betingelse; ansikte och röster samtidigt. Precis så som Cook och Wilding (2007) och Stevenage et al. (2011) testade men där eventuell expertis fungerar som ytterligare en variabel.

Referenser

- Borgström, C. (2007). *Röstigenkänning och Ansiktsgenkänning Kan yrkeserfarenhet påverka prestation och konfidensbedömningar?* C-uppsats, Högskolan Kristianstad, Institutionen för Beteendevetenskap.
- Chin, J. M., & Schooler, J. W. (2008): Why do words hurt? Content, process, and criterion shift accounts of verbal overshadowing, *European Journal of Cognitive Psychology*, 20(3), 396-413.
- Christianson, S. A., Karlsson I., & Persson, L. G. W. (1998) Police Personnel as Eyewitnesses to a Violent Crime. *Legal and Criminological Psychology*, 3(1) 59–72.
- Clifford, B. R., & Bull, R. (1978). *The psychology of person identification*. London: Routledge Kegan Paul Ltd.
- Clifford, B. R. (1980). Voice Identification by Humans Listeners: On Earwitness Reliability. *Law and Human Behavior*, 4(4), 373-394.
- Coob, N., Lawrence, D., & Nelson, N. (1979). Report on blind subjects' tactile and auditory recognition for environmental stimuli. *Perceptual and Motor Skills*, 48(2), 363-366.
- Cook, S., & Wilding, J. (2007). Earwitness testimony 2: Voices, faces and context. *Applied cognitive psychology*, 11(6), 527-541.
- Dodson, C. S., Johnson, M. K., & Schooler, J.W. 1997, The verbal overshadowing effect: Why descriptions impair face recognition. *Memory & Cognition*, 25(2), 129-139.
- Engstrand, O. (2004). *Fonetikens grunder*. Lund: Studentlitteratur.
- Granhag, P. A. 2001. *Vittnespsykologi*. Lund: Studentlitteratur.
- Granhag, P. A. & Christianson, S. Å. (Red.). (2008). *Handbok i rättspsykologi*. Stockholm: Liber.
- Hollien, H. (2002). *Forensic Voice Identification*. San Diego: Academic Press.
- McGuire, J. 2004. *Understanding psychology and crime*. Berkshire: Open University Press.
- Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language*, 35(2), 231-245.
- Mitchell, H. F., & MacDonald, R. A. R. (2011). Remembering, Recognizing and Describing Singers' Sound Identities. *Journal of New Music Research*, 40(1), 75-80.
- Mullenix, J.W., Stern, S. S., Grounds, S., Kalas, R., Flaherty, M., Kowalok, S., et al. (2010). Earwitness Memory: Distortions for Voice Pitch and Speaking Rate. *Applied Cognitive Psychology*, 24(4), 513-526.

- Olsson, N., Juslin, P., & Winman, A. (1998). Realism of confidence in earwitness versus eyewitness identification. *Journal of experimental psychology:Applied*, 4(2), 101-118.
- Philippon, A. C., Cherryman, J., Bull, R., & Vrij, A. (2007). Lay People's and Police Officers' Attitudes Towards the Usefulness of Perpetrator Voice Identification. *Applied Cognitive Psychology*, 21(1), 103-115.
- Räma, P., & Courtney, S. M. (2005). Functional topography of working memory for face or voice identity. *NeuroImage* 24(1), 224-234.
- Read, D., & Craik, F. I. M. (1995). Earwitness Identification: Some Influences on Voice Recognition. *Journal of Experimental Psychology: Applied*, 1(1), 6-18.
- Rikspolisstyrelsen. (2005). *RPS Rapport 2005:2 Vittneskonfrontation*. Stockholm: Rikspolisstyrelsen.
- Schacter, D. L. (2001). *The seven sins of memory: How the mind forgets and Remembers*. New York: Houghton Mifflin Company.
- Schooler, J.W., & Engstler-Schooler, T.Y. (1990). Verbal overshadowing of visual memories: Some things are better unsaid. *Cognitive Psychology*, 22(1), 36-71.
- Solan, M. & Tiersma, P. M. (2003). Falling on deaf ears. *Legal Affairs*. November/December.
- Stevenage, S.V., Howland, A., & Tippelt, A. (2011). Interference in eyewitness and earwitness recognition. *Applied cognitive psychology*, 25(1), 112-118.
- Vanags, T., Carroll, M., & Perfect, T.J., (2005). Verbal Overshadowing: A Sound Theory in Voice Recognition? *Applied Cognitive Psychology*, 19(9), 1127-1144.
- Werkelin, P., & Svanberg, K. (2006). *Klockan 10.31 på morgonen I Khao Lak*. Stockholm: Bonnier fakta
- Winograd, E., Kerr, N. H., & Spence, M. J. (1984) Voice Recognition: Effects of Orienting Task, and a Test of Blind versus Sighted Listeners. *The American journal of psychology*, 97 (1), 57-70.
- Yarmey, A. D. (1986). Verbal, Visual and Voice Identification of a Rape Suspect Under Different Levels of Illumination. *Journal of Applied Psychology*, 71(3), 363-370.
- Yarmey, A. D. (1995). Earwitness speaker identification. *Public Policy and Law Psychology*, 1(4), 792-816.
- Yarmey, A. D., Yarmey, A. L., Yarmey, M. J., & Parliament, L. (2001). Commonsense Beliefs and the Identification of Familiar Voices. *Applied Cognitive Psychology*, 15(3), 283-299.